Gag_AF110965_BW_mod

ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGGCAAGCTGGACGCCTGGGAGCGCATCCGCC
TGCGCCCCGGCGGCAAGAAGTGCTACATGATGAAGCACCTGGTGTGGGCCAGCCGCGAGCT
GGAGAAGTTCGCCCTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATCATC
CGCCAGCTGCACCCCGCCCTGCAGACCGGCAGCGAGGAGCTGAAGAGCCTGTTCAACACCG
TGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTCCGCGACACCAAGGAGGCCCTGGA
CAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGATCCAGCAGGCCGAGGCCGCCGAC
AAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACC
AGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAGGCCTTCAG
CCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAGGACCTGAAC
ACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAGGACACCATCA
ACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCCGTGCACGCCGGCCCCATCGCCCCCGG
CCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGCACCACCAGCACCCTGCAGGAGCAG
ATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTACAAGCGGTGGATCA
TCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATCCTGGACATCAAGCA
GGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCCTGCGCGCCGAG
CAGAGCACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTGGTGCAGAACGCCAACC
CCGACTGCAAGACCATCCTGCGCGCTCTCGGCCCCGGCGCCAGCCTGGAGGAGATGATGAC
CGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTGCTGGCCGAGGCGATGAGC
CAGGCCAACACCAGCGTGATGATGCAGAAGAGCAACTTCAAGGGCCCCCGGCGCATCGTCA
AGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAACTGCCGCGCCCCCCGCAAGAA
GGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCC
AACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGCCCCGGCAACTTCCTGCAGAGCC
GCCCCGAGCCCACCGCCCCCCCCGCCGAGAGCTTCCGCTTCGAGGAGACCACCCCCGGCCA
GAAGCAGGAGAGCAAGGACCGCGAGACCCTGACCAGCCTGAAGAGCCTGTTCGGCAACGAC
CCCCTGAGCCAGTAA

Figure 1

Gag_AF110967_BW_mod

ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGAGAAGCTGGACAAGTGGGAGAAGATCCGCC
TGCGCCCCGGCGGCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAGCT
GGAGGGCTTCGCCCTGAACCCCGGCCTGCTGGAGACCGCCGAGGGCTGCAAGCAGATCATG
AAGCAGCTGCAGCCCGCCCTGCAGACCGGCACCGAGGAGCTGCGCAGCCTGTACAACACCG
TGGCCACCCTGTACTGCGTGCACGCCGGCATCGAGGTCCGCGACACCAAGGAGGCCCTGGA
CAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAAGACCCAGCAGGCCAAGGAGGCCGAC
GGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCAGG
CCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAGGCCTTCAGCCC
CGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAGGACCTGAACACG
ATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAGGACACCATCAACG
AGGAGGCCGCCGAGTGGGACCGCCTGCACCCCGTGCAGGCCGGCCCCGTGGCCCCCGGCCA
GATGCGCGACCCCCGCGGCAGCGACATCGCCGGCGCCACCAGCACCCTGCAGGAGCAGATC
GCCTGGATGACCAGCAACCCCCCCGTGCCCGTGGGCGACATCTACAAGCGGTGGATCATCC
TGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATCCTGGACATCCGCCAGGG
CCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCCTGCGCGCCGAGCAG
GCCACCCAGGACGTGAAGAACTGGATGACCGAGACCCTGCTGGTGCAGAACGCCAACCCCG
ACTGCAAGACCATCCTGCGCGCTCTCGGCCCCGGCGCCACCCTGGAGGAGATGATGACCGC
CTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTGCTGGCCGAGGCGATGAGCCAG
GCCAACAGCGTGAACATCATGATGCAGAAGAGCAACTTCAAGGGCCCCCGGCGCAACGTCA
AGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAAGAACTGCCGCGCCCCCCGCAAGAA
GGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCC
AACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGCCCCGGCAACTTCCTGCAGAACC
GCAGCGAGCCCGCCGCCCCCACCGTGCCCACCGCCCCCCCCGCCGAGAGCTTCCGCTTCGA
GGAGACCACCCCCGCCCCCAAGCAGGAGCCCAAGGACCGCGAGCCCTACCGCGAGCCCCTG
ACCGCCCTGCGCAGCCTGTTCGGCAGCGGCCCCCTGAGCCAGTAA

Figure 2

Fig. 3

Env_AF110968_C_BW_opt

--> signal peptide (1-81)
ATGCGCGTGATGGGCATCCTGAAGAACTACCAGCAGTGGTGGATGTGGGGCATCCTGGGCTTCTGGATGCTGATCA
\/--> gp120/140/160 (82)
TCAGCAGCGTGGTGGGCAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAAGACCACCCT

GTTCTGCACCAGCGACGCCAAGGCCTACGAGACCGAGGTGCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACC

GACCCCAACCCCCAGGAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACGACATGGTGGACC

AGATGCACGAGGACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGAC

CCTGAAGTGCCGCAACGTGAACGCCACCAACAACATCAACAGCATGATCGACAACAGCAACAAGGGCGAGATGAAG

AACTGCAGCTTCAACGTGACCACCGAGCTGCGCGACCGCAAGCAGGAGGTGCACGCCCTGTTCTACCGCCTGGACG

TGGTGCCCCTGCAGGGCAACAACAGCAACGAGTACCGCCTGATCAACTGCAACACCAGCGCCATCACCCAGGCCTG

CCCCAAGGTGAGCTTCGACCCCATCCCCATCCACTACTGCACCCCCGCCGGCTACGCCATCCTGAAGTGCAACAAC

CAGACCTTCAACGGCACCGGCCCCTGCAACAACGTGAGCAGCGTGCAGTGCGCCCACGGCATCAAGCCCGTGGTGA

GCACCCAGCTGCTGCTGAACGGCAGCCTGGCCAAGGGCGAGATCATCATCCGCAGCGAGAACCTGGCCAACAACGC

CAAGATCATCATCGTGCAGCTGAACAAGCCCGTGAAGATCGTGTGCGTGCGCCCCAACAACAACACCCGCAAGAGC

GTGCGCATCGGCCCCGGCCAGACCTTCTACGCCACCGGCGAGATCATCGGCGACATCCGCCAGGCCTACTGCATCA

TCAACAAGACCGAGTGGAACAGCACCCTGCAGGGCGTGAGCAAGAAGCTGGAGGAGCACTTCAGCAAGAAGGCCAT

CAAGTTCGAGCCCAGCAGCGGCGGCGACCTGGAGATCACCACCCACAGCTTCAACTGCCGCGGCGAGTTCTTCTAC

TGCGACACCAGCCAGCTGTTCAACAGCACCTACAGCCCCAGCTTCAACGGCACCGAGAACAAGCTGAACGGCACCA

TCACCATCACCTGCCGCATCAAGCAGATCATCAACATGTGGCAGAAGGTGGGCCGCGCCATGTACGCCCCCCCCAT

CGCCGGCAACCTGACCTGCGAGAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGACCGGCCCCAAC

GACACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGACAACTGGCGCAACGAGCTGTACAAGTACAAGGTGG
gp120(1512)<--\/-->(1513)gp41
TGGAGATCAAGCCCCTGGGCGTGGCCCCCACCGAGGCCAAGCGCCGCGTGGTGGAGCGCGAGAAGCGCGGCCGTGGG

CATCGGCGCCGTGTTCCTGGGCTTCCTGGGCGCCGCCGGCAGCACCATGGGCGCCGCCAGCATCACCCTGACCGTG

CAGGCCCGCCTGCTGCTGAGCGGCATCGTGCAGCAGCAGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACC

TGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGACCCGCATCCTGGCCGTGGAGCGCTACCTGAAGGACCA

GCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACAGCAGCTGGAGC

AACCGCAGCCACGACGAGATCTGGGACAACATGACCTGGATGCAGTGGGACCGCGAGATCAACAACTACACCGACA

CCATCTACCGCCTGCTGGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGAAGGACCTGCTGGCCCTGGACAGCTG
gp140(2025)<--\/
GCAGAACCTGTGGAACTGGTTCAGCATCACCAACTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGC

CTGATCGGCCTGCGCATCATCTTCGCCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCCTGCCCT

TCCAGACCCTGACCCCCAACCCCCGCGAGCCCGACCGCCTGGGCCGCATCGAGGAGGAGGGCGGCGAGCAGGACCG

CGGCCGCAGCATCCGCCTGGTGAGCGGCTTCCTGGCCCTGGCCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGC

TACCACCGCCTGCGCGACTTCATCCTGATCGCCGCCCGCGTGCTGGAGCTGCTGGGCCAGCGCGGCTGGGAGGCCC

TGAAGTACCTGGGCAGCCTGGTGCAGTACTGGGGCCTGGAGCTGAAGAAGAGCGCCATCAGCCTGCTGGACACCAT

CGCCATCGCCGTGGCCGAGGGCACCGACCGCATCATCGAGTTCATCCAGCGCATCTGCCGCGCCATCCGCAACATC
gp160, gp41(2547)<--
CCCCGCCGCATCCGCCAGGGCTTCGAGGCCGCCCTGCAGTAA

Fig. 4

Env_AF110975_C_BW_opt

```
--> signal peptide (1-72)                                              \/-->
ATGCGCGTGCGCGGCATCCTGCGCAGCTGGCAGCAGTGGTGGATCTGGGGCATCCTGGGCTTCTGGATCTGCAGCG
gp120/140/160 (72)
GCCTGGGCAACCTGTGGGTGACCGTGTACGACGGCGTGCCCGTGTGGCGCGAGGCCAGCACCACCCTGTTCTGCGC

CAGCGACGCCAAGGCCTACGAGAAGGAGGTGCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAAC

CCCCAGGAGATCGAGCTGGACAACGTGACCGAGAACTTCAACATGTGGAAGAACGACATGGTGGACCAGATGCACG

AGGACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCCGCGTGAAGCTGACCCCCCTGTGCGTGACCCTGAAGTG

CACCAACTACAGCACCAACTACAGCAACACCATGAACGCCACCAGCTACAACAACAACACCACCGAGGAGATCAAG

AACTGCACCTTCAACATGACCACCGAGCTGCGCGACAAGAAGCAGCAGGTGTACGCCCTGTTCTACAAGCTGGACA

TCGTGCCCCTGAACAGCAACAGCAGCGAGTACCGCCTGATCAACTGCAACACCAGCGCCATCACCCAGGCCTGCCC

CAAGGTGAGCTTCGACCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTACGCCATCCTGAAGTGCAAGAACAAC

ACCAGCAACGGCACCGGCCCCTGCCAGAACGTGAGCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGTGAGCA

CCCCCCTGCTGCTGAACGGCAGCCTGGCCGAGGGCGGCGAGATCATCATCCGCAGCAAGAACCTGAGCAACAACGC

CTACACCATCATCGTGCACCTGAACGACAGCGTGGAGATCGTGTGCACCCGCCCCAACAACAACACCCGCAAGGGC

ATCCGCATCGGCCCCGGCCAGACCTTCTACGCCACCGAGAACATCATCGGCGACATCCGCCAGGCCCACTGCAACA

TCAGCGCCGGCGAGTGGAACAAGGCCGTGCAGCGCGTGAGCGCCAAGCTGCGCGAGCACTTCCCCAACAAGACCAT

CGAGTTCCAGCCCAGCAGCGGCGGCGACCTGGAGATCACCACCCACAGCTTCAACTGCCGCGGCGAGTTCTTCTAC

TGCAACACCAGCAAGCTGTTCAACAGCAGCTACAACGGCACCAGCTACCGCGGCACCGAGAGCAACAGCAGCATCA

TCACCCTGCCCTGCCGCATCAAGCAGATCATCGACATGTGGCAGAAGGTGGGCCGCGCCATCTACGCCCCCCCCCAT

CGAGGGCAACATCACCTGCAGCAGCAGCATCACCGGCCTGCTGCTGGCCCGCGACGGCGGCCTGGACAACATCACC

ACCGAGATCTTCCGCCCCCAGGGCGGCGACATGAAGGACAACTGGCGCAACGAGCTGTACAAGTACAAGGTGGTGG
                                                     gp120(1509)<--\/-->(1510)gp41
AGATCAAGCCCCTGGGCGTGGCCCCCACCGAGGCCAAGCGCCGCGTGGTGGAGCGCGAGAAGCGCGCCGTGGGCAT

CGGCGCCGTGATCTTCGGCTTCCTGGGCGCCGCCGGCAGCAACATGGGCGCCGCCAGCATCACCCTGACCGCCCAG

GCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGAGCAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACATGC

TGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCATCGAGCGCTACCTGAAGGACCAGCA

GCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCACCGTGCCCTGGAACAGCAGCTGGAGCAAC

AAGACCCAGGGCGAGATCTGGGAGAACATGACCTGGATGCAGTGGGACAAGGAGATCAGCAACTACACCGGCATCA

TCTACCGCCTGCTGGAGGAGAGCCAGAACCAGCAGGAGCAGAACGAGAAGGACCTGCTGGCCCTGGACAGCCGCAA
                             gp140(2022)<--\/
CAACCTGTGGAGCTGGTTCAACATCAGCAACTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTG

ATCGGCCTGCGCATCATCTTCGCCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCC

AGACCCTGACCCCCAACCCCCGCGGCCTGGACCGCCTGGGCCGCATCGAGGAGGAGGGCGGCGAGCAGGACCGCGA

CCGCAGCATCCGCCTGGTGCAGGGCTTCCTGGCCCTGGCCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTAC

CACCGCCTGCGCGACCTGATCCTGGTGACCGCCCGCGTGGTGGAGCTGCTGGGCCGCAGCAGCCCCCGCGGCCTGC

AGCGCGGCTGGGAGGCCCTGAAGTACCTGGGCAGCCTGGTGCAGTACTGGGGCCTGGAGCTGAAGAAGAGCGCCAC

CAGCCTGCTGGACAGCATCGCCATCGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGATCCAGCGCATCTAC
                                 gp160, gp41(2565)<--\
CGCGCCTTCTGCAACATCCCCCGCCGCGTGCGCCAGGGCTTCGAGGCCGCCCTGCAGTAA
```

Gag_AF110965_BW_opt

ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGGCAAGCTGGACGCCTGGGAGCGCATCCGCCTGCGCCCCGG

CGGCAAGAAGTGCTACATGATGAAGCACCTGGTGTGGGCCAGCCGCGAGCTGGAGAAGTTCGCCCTGAACC

CCGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATCATCCGCCAGCTGCACCCCGCCCTGCAGACCGGC

AGCGAGGAGCTGAAGAGCCTGTTCAACACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTGCG

CGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGATCCAGCAGGCCG

AGGCCGCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCAC

CAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAGGCCTTCAGCCCCGAGGT

GATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAGGACCTGAACACCATGCTGAACACCGTGG

GCGGCCACCAGGCCGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTG

CACCCCGTGCACGCCGGCCCCATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGCAC

CACCAGCACCCTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTACA

AGCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACAGCCCCGTGAGCATCCTGGACATCAAG

CAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCCTGCGCGCCGAGCAGAGCAC

CCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCC

TGCGCGCCCTGGGCCCCGGCGCCAGCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGC

CACAAGGCCCGCGTGCTGGCCGAGGCCATGAGCCAGGCCAACACCAGCGTGATGATGCAGAAGAGCAACTT

CAAGGGCCCCCGCCGCATCGTGAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAACTGCCGCG

CCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAG

GCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGCCCCGGCAACTTCCTGCAGAGCCGCCCCGA

GCCCACCGCCCCCCCCGCCGAGAGCTTCCGCTTCGAGGAGACCACCCCCGGCCAGAAGCAGGAGAGCAAGG

ACCGCGAGACCCTGACCAGCCTGAAGAGCCTGTTCGGCAACGACCCCCTGAGCCAGTAA

Figure 5

Gsg_AF110967_BW_

ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGAGAAGCTGGACAAGTGGGAXAGATCCGCCTGCGCCCCGG

CGGCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAGCTGGAGGGCTTCGCCCTGAACC

CCGGCCTGCTGGAGACCGCCGAGGGCTGCAAGCAGATCATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGC

ACCGAGGAGCTGCGCAGCCTGTACAACACCGTGGCCACCCTGTACTGCGTGCACGCCGGCATCGAGGT[C]CG

CGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGCAGAACAAG[AG]CCAGCAGAAGACCCAGCAGGCCA

AGGAGGCCGACGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCAG

GCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAGGCCTTCAGCCCCGAGGTGAT

CCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCCAGGACCTGAACAG[CATGC]TGAACACCGTGGGCG

GCCACCAGGCCGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCTGCAC

CCCGTGCAGGCCGGCCCCGTGGCCCCCGGCCAGATGCGCGACCCCCGCGGCAGCGACATCGCCGGCGCCAC

CAGCACCCTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCGTGCCCGTGGGCGACATCTACAAGC

[GC]TGGATCATCCTGGGCCTGAACAAGATCGTGCG[C]ATGTACAGCCCCGTGAGCATCCTGGACATCCGCCAG

GGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCCTGCGCGCGCCGAGCAGGCCACCCA

GGACGTGAAGAACTGGATGACCGAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGC

GCG[CCTG]GGCCCCGGCGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCAC

AAGGCCCGCGTGCTGGCCGAGGG[C]ATGAGCCAGGCCAACAGCGTGAACATCATGATGCAGAAGAGCAACTT

CAAGGGCCCCCG[C]GCAACGT[C]AAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAAGAACTGCCGCG

CCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAG

GCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGCCCCGGCAACTTCCTGCAGAACCGCAGCGA

GCCCGCCGCCCCCACCGTGCCCACCGCCCCCCCCGCCGAGAGCTTCCGCTTCGAGGAGACCACCCCCGCCC

CCAAGCAGGAGCCCAAGGACCGCGAGCCCTACCGCGAGCCCCTGACCGCCCTGCGCAGCCTGTTCGGCAGC

GGCCCCCTGAGCCAGTAA

Figure 6